

## コンピュータ基礎実習上級 #11 画像ファイルと文字のフォーマット

一般教育研究センター 安田豊

## 画像ファイルのフォーマット

- ペイントで作った画像が表示されない
  - ブラウザによっては表示される
  - 何故か？
- 画像にはフォーマット(書式)がある
  - 書式 = データの書き方
  - データ化の方法からはじめないと判らない

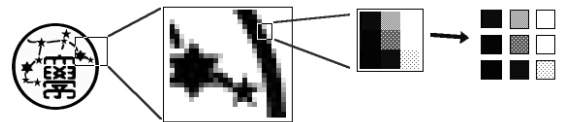
## 先に結論だけ

- 幾つかの画像フォーマットがある
  - それぞれ美点と欠点がある
  - 使い分けられて共存

jpg	JPEG	写真など多くの色を扱うのに便利。圧縮が可能。
bmp	ビットマップ	Microsoftが作ったフォーマット。圧縮していない。
gif	GIF	CompuServe社が作成。小さな画像に適する。
png	PNG	UNISYS/パテント問題に対応するために現在利用が広がっている。

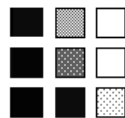
## デジタルデータの表現

- 絵は画素(Pixel)に分解
  - 各画素ごとに数値化
  - 1-3-5, 1-2-5, 1-1-4 のごとし



## フォーマット

- 数え方を違えると全く違う数字列に
  - 1-3-5, 1-2-5, 1-1-4
  - 1-1-1, 3-2-1, 5-5, 4
- 数値化ルールと一致する復号をしないと違う結果になる



## コード化ルール

- デジタルデータは常に数値化したルールとセットでなければ扱えない
  - 数値化をエンコード
  - 復号をデコードと呼ぶ
- コード化ルールがデータのフォーマットを決める
  - 「こうする」と決めた結果「こう書いた」
  - 画像データは複数のコード化ルールが普及している世界

## 文字コード

- 文字も同様にデジタル化
  - A=1,B=2...Z=26
  - a=27,b=28...z=52 (あくまで例)
  - 漢字も数千文字をこうやって数値化
- 同じコード化ルールでなければ表示させることができない
  - いわゆる文字化けを起こす

## WWWと文字コード

- 国内では主として三つの文字コードが利用されている(\*)
  - ISO-2022-JP, EUC-JP, Shift-JIS
  - ( \* 現実はずっと複雑だが今回は説明せず )
- HTML作成時に選択可能
  - テキストエディタに選択機能がある
- METAタグによる明示
  - <META HTTP-EQUIV="Content-Type"
  - CONTENT="text/html;charset=ISO-2022-JP">

## まとめ

- デジタルデータは
  - コード化ルールとセット
  - encode / decode に不整合があると
    - 絵がおかしくなったり
    - 文字が別の文字になったりする
- WWW制作の場面では
  - 画像ファイルは拡張子に注意
  - 文字コードはできればMETAタグで明示